ACM Digital Library
search

Search the Digital Library

Search Articles:

Terms: ranking documents → 256 articles found

◉ all words ○ any words ○ exact phrase ○ subject ○ expression (☐ stem)

In Fields:
- ☑ Title (50,699)
- ☑ Full-Text (40,518)
- ☐ Abstract (12,474)
- ☐ Reviews (2,602)
- ☐ Index Terms (38,489)

(Number of articles)

Authors:

◉ all names ○ any name ○ expression (☐ soundex)

Limit Your Search To:

Publication: All Journals and Proceedings

Published Since: Month Year

Published Before: January 1997

search

[Help]

Annual ACM Conference on Research and Development in Information Retrieval
Proceedings of the Fifteenth Annual International ACM SIGIR conference on Research and development in information retrieval
June 21 - 24, 1992, Copenhagen Denmark

Integration of probabilistic fact and text retrieval

Page 211

Norbert Fuhr

ABSTRACT

In this paper, a model for combining text and fact retrieval is described. A query is a set of conditions, where a single condition is either a text or fact condition. Fact conditions can be interpreted as being vague, thus leading to nonbinary weights for fact conditions with respect to database objects. For text conditions, we use descriptions of the occurence of terms in documents instead of precomputed indexing weights, thus treating terms similar to attributes. Probabilistic indexing weights for conditions are computed by introducing the notion of correctness (or acceptability) of a condition w.r.t. an object. These indexing weights are used in retrieval for a probabilistic ranking of objects based on the retrieval for a probabilistic ranking of objects based on the retrieval-with-probabilistic-indexing (RPI) model, for which a new derivation is given here.

INDEX TERMS

Categories and Subject Descriptors:

Information Systems -Information Storage and Retrieval - Content Analysis

and Indexing (H.3.1): **Indexing methods**; Information Systems -Information Storage and Retrieval - Information Search and Retrieval (H.3.3): **Retrieval models**; Computing Methodologies -Document and Text Processing - Document and Text Editing (I.7.1); Information Systems -Information Interfaces and Presentation - User Interfaces (H.5.2);

General Terms:
Algorithms, Languages

Annual ACM Conference on Research and Development in Information Retrieval

Proceedings of the sixteenth annual international ACM SIGIR conference on Research and Development in Information Retrieval June 27 - July 1, 1993, Pittsburgh, PA USA

A probabilistic relational model for the integration of IR and databases

Page 309

Norbert Fuhr

ABSTRACT

In this paper, a probabilistic relational model is presented which combines relational algebra with probabilistic retrieval. Based on certain independence assumptions, the operators of the relational algebra are redefined such that the probabilistic algebra is a generalization of the standard relational algebra. Furthermore, a special join operator implementing probabilistic retrieval is proposed. When applied to typical document databases, queries can not only ask for documents, but for any kind of object in the database. In addition, an implicit ranking of these objects is provided in case the query relates to probabilistic indexing or uses the probabilistic join operator. The proposed algebra is intended as a standard interface to combined database and IR systems, as a basis for implementing user-friendly interfaces.

INDEX TERMS

Categories and Subject Descriptors:

Information Systems -Information Storage and Retrieval - Information Search

and Retrieval (H.3.3): **Retrieval models**; Information Systems -Database Management - Languages (H.2.3): **SQL**; Information Systems -Database Management - Logical Design (H.2.1): **Data models**; Computing Methodologies -Symbolic and Algebraic Manipulation - General (I.1.0);

General Terms:
Algorithms, Design, Languages

Annual ACM Conference on Research and Development in Information Retrieval

Proceedings of the sixteenth annual international ACM SIGIR conference on Research and Development in Information Retrieval June 27 - July 1, 1993, Pittsburgh, PA USA

A user-centred evaluation of ranking algorithms for interactive query expansion

Page 146

Efthimis N.Efthimiadis

ABSTRACT

The evaluation of 6 ranking algorithms for the ranking of terms for query expansion is discussed within the context of an investigation of interactive query expansion and relevance feedback in a real operational environment. The yardstick for the evaluation was provided by the user relevance judgements on the lists of the candidate terms for query expansion. The evaluation focuses on the similarities in the performance of the different algorithms and how the algorithms with similar performance treat terms.

INDEX TERMS

General Terms:
Algorithms, Performance, Theory

Categories and Subject Descriptors:
Information Systems -Information Storage and Retrieval - Information Search and Retrieval (H.3.3): **Query formulation**; Information Systems -Information

Storage and Retrieval - Content Analysis and Indexing (H.3.1): **Abstracting methods**;

ACM Symposium on Parallel Algorithms and Architectures
Proceedings of the 6th annual ACM symposium on Parallel algorithms and architectures
June 27 - 29, 1994, Cape May, NJ USA

List ranking and list scan on the Cray C-90

Pages 104-113

Margaret Reid-Miller

ABSTRACT

List ranking and list scan are two primitive operations used in many parallel algorithms that use list, trees, and graph data structures. But vectorizing and parallelizing list ranking is a challenge because it is highly communication intensive and dynamic. In addition, the serial algorithm is very simple and has very small constants. In order to compete, a parallel algorithm must also be simple and have small constants. A parallel algorithm due to Wyllie is such an algorithm, but it is not work efficient--its performance degrades for longer and longer linked lists. In contrast, work efficient PRAM algorithms developed to date have very large constants. It does not achieve O(log *n*) running time, but we contend that work efficiency and small constants is more important, given that vector and multiprocessor machines are used for problems that are much larger than the number of processors and, therefore, the O(log *n*) running time, but we contend that work efficiency and small constants is more important, given that vector and multiprocessor machines are used for problems that are much larger than the number of processors and, therefore, the O(log *n*) time is never achieved in practice. In particular, to the best of our knowledge, our implementation of list ranking and list scan on the CRAY C-90 is the fastest implementation to date. In addition, it is the first implementation of which we are aware that outperforms fast workstations. The success of our algorithm is due to its relatively large grain size and simplicity of the inner loops, and the success of the implementation is due to pipelining reads and writes through vectorization to hide latency, minimizing load balancing by

deriving equations for predicting and optimizing performance, and avoiding conditional tests except when load balancing.

INDEX TERMS

Categories and Subject Descriptors:
Theory of Computation -Analysis of Algorithms and Problem Complexity - Nonnumerical Algorithms and Problems (F.2.2); Computer Systems Organization -Processor Architectures - Multiple Data Stream Architectures (Multiprocessors) (C.1.2); Theory of Computation -Computation by Abstract Devices - Modes of Computation (F.1.2): **Parallelism and concurrency**;

General Terms:
Algorithms, Measurement, Performance, Theory

REVIEWS

From Computing Reviews
William Fennell Smyth

Given a list L of n elements $x_1,x_2,,x_n$, the list scan problem requires that, at each position i of L , the sum $x_1+x_2++x_i$ be formed, where "+" is some binary associative operator. The list ranking problem is the special case of list scan that arises when "+" signifies ordinary addition and every $x_i=1$. List scan occurs frequently as a subproblem in many parallel combinatorial algorithms.

This paper describes a new list scan algorithm and gives its implementation on the Cray C-90 vector multi^processor. The new algorithm is both work efficient (that is, it executes in Qn time) and fast (that is, the constants of proportionality are small), and for large n , its execution time on the C-90 is an order of magnitude faster than that of other known algorithms. The main idea of the new algorithm is to break up L into m sublists, where usually nmp , if p is the number of processors; each processor then deals with m/p sublists. To compensate for variation in the lengths of the sublists, periodic load balancing is carried out: unprocessed elements in long sublists are packed together into contiguous locations. The author points out that, since the C-90 can be thought of as approximating an exclusive read exclusive write parallel random access machine (EREW PRAM), the new algorithm may provide a basis for the efficient execution of known PRAM algorithms that depend on list scan for their execution.

The paper is interesting and well written, but it suffers from numerous syntactical and grammatical anomalies that would certainly have been eliminated by thorough copyediting and proofreading.